

HYPOTHESES CONCERNING THE PHONETIC AND FUNCTIONAL ORIGINS OF TONE DISPLACEMENT IN ZULU

Daniel Silverman
Department of Linguistics
Beckman Institute for Advanced Science and Technology
University of Illinois

High tones in Zulu are displaced rightward from their vowel of origin in the context of preceding “depressor” consonants, a segment type traditionally considered to possess breathy phonation. In this paper, I suggest that physical properties of the speech mechanism—phonetic factors—may have initially induced the apparent rightward “unhinging” of high tones in the context of preceding depressors, and that independent functional factors may have induced the conventionalization of tone displacement. As tones were less likely to neutralize upon displacement, displaced tokens were more often perceived unambiguously, and, hence, were more likely to be reproduced.

1. Introduction

Every linguistic sound system is the product of its own unique history. As phonetic, cognitive, and social forces interact over the generations, each language, each dialect, and ultimately each idiolect, maintains uniqueness. Of course, documenting completely the interactions among these forces is a largely impossible task. But if we investigate a small corner of the sound system across a large population of speakers, we might begin to disentangle the web of forces that are responsible for its patterning. In this paper, I present hypotheses about some of the phonetic, cognitive, and social forces that may have given rise to a sound pattern found in Zulu and related Nguni languages: high tones are displaced rightward in the context of preceding “depressor” consonants (see, for example, Beach

[1924], Doke [1926], Lanham [1958], Cope [1960], [1966], Schachter [1976], Rycroft [1980], Khumalo [1981], Traill et al. [1987], Traill & Jackson [1988]). Many researchers report that depressor consonants (“D”) are accompanied by breathy phonation on the following vowel [Tucker 1949, Cope 1960, Rycroft 1980], as well as extreme pitch lowering at the consonantal offset. Following depressor consonants, high-tones on short vowels are displaced from their vowel of origin to a following vowel ($DV\uparrow CV\downarrow \rightarrow DV\downarrow CV\uparrow$; where D=depressor, C=non-depressor). Moreover, tone displacement is blocked if a depressor immediately follows ($DV\uparrow DV\downarrow \rightarrow DV\uparrow DV\downarrow$). I consider research results from a number of disciplines—phonology, phonetics, psychology—and suggest ways in which the general principles which seem to underlie these results may have interacted over time to culminate in the Zulu sound pattern in evidence today. I argue that the diachronic “trigger” of high tone displacement in the context of preceding depressor consonants may have been phonetic in origin. However, there may be functional reasons why the pattern has been conventionalized in the present-day manner.

While readers may have expertise in one or more of the areas considered herein, I hope to impart a general understanding of how the various forces that seem to be at work might have interacted over time. Indeed, phoneticians may be standing at the ready to provide compelling phonetic motivation for the patterns in evidence; variationists might offer a statistical model of the sound pattern; phonologists might provide a formal model of the broadly categorical elements of the pattern; cognitive scientists may apply general learning theories to the acquisition of the pattern; historical and comparative linguists might document the sound change over time and space. I hope to show that research in all of these areas is necessary—if not necessarily sufficient—to achieve an explanatorily satisfying account of sound patterns, both in Zulu and elsewhere. The goal, then, is not to provide exhaustive discussions of the various forces argued to be at work here. Rather, I aim to show how experts from a number of fields might combine their talents to arrive at a satisfying explanation for the sound pattern in question; simple, general, and broadly applicable hypotheses are a hallmark of compelling scientific investigations.

Given that experts in these fields employ rather different research methodologies, this paper is structured to give voice to each, in turn. In section 2, I provide a literature review of linguistic descriptions of the Zulu tone displacement pattern. In Section 3, I exemplify the pattern, discussing in detail the phonological environments in which tone displacement is observed, and—equally important—the environments in which tone displacement is not observed. In Section 4, I consider the phonetic motivation for tone displacement, offering a tutorial on the

aerodynamic and articulatory constraints that may have set the pattern into diachronic motion:

- Depressor consonants have been characterized as phonetically and/or historically breathy-voiced.
- Obstruent voicing in general, and obstruent breathy-voicing in particular, is associated with pitch lowering at consonant release.
- Pitch rises take longer to implement than do pitch falls.
- In the context of a depressor consonant, a following high tone may be achieved only after the following consonantal gesture has been completed, culminating in an apparent displacement of the high tone.

I further report on a recent instrumental study [Russell 2000] which numerically documents the relevant pitch perturbation patterns.

In Section 5, I depart from phonetic considerations and instead focus on functional forces which may have led to the present-day convention of tone displacement:

- If tone displacement were not to take place, it is possible that these high tones would not be accurately conveyed to listeners.
- As these high tones are contrastive in Zulu, non-displacement might run the risk of neutralizing a significant number of lexical distinctions.
- Displaced high tones were therefore more likely to maintain lexical contrasts than non-displaced high tones.

In Section 6, I discuss the possible diachronic interactions of the phonetic and functional influences on the pattern, in the context of certain general cognitive abilities documented in humans and many lower animals as well:

- There is inherent variability in speech production; both displaced and non-displaced tokens are among the possible variants.
- Learners come to largely reproduce the nuances of variation they perceive their elders engaging in, despite the fact that variants with displaced tones are more successful at keeping contrastive elements distinct. That is, the variability engaged in by elders will be largely matched by learners.
- Ambiguous tokens will sometimes be impossible to categorize, and, hence, will not be added to the pool of tokens over which variability is calculated.
- Due to the greater likelihood of unambiguous perception of displaced variants, learners' calculated variability may differ slightly from their elders' in that the variants which contrast more sharply with oppositions will more often be perceived correctly, hence, in turn, be more likely produced.
- The variability inherent in speech production may be the fodder for these sorts of sounds changes; the more distinct the variant from an acoustically

similar contrastive value, the more likely the system will wend towards this variant.

I suggest then, that tone displacement had a phonetic trigger, but that a listener-based functional reaction led to a conventionalization of the pattern.

Finally, in Section 7, I briefly discuss other patterns from a variety of unrelated languages that seem amenable to a similar explanation.

2. Overview of previous literature on depressors and their effects on pitch

Cope [1966] reports that the consonants listed in (1) belong to the depressor class. He claims that depressors are characterized by their “heaviness” (p.113), that is, by a combination of breath and voice, although not all researchers agree that this is the proper phonetic characterization of the class. Cope further reports that Zulu possesses three contrastive tone patterns, high, low, as well as a high-low toneme cluster.

(1) Zulu depressors

voiced aspirated stops	b ^h	d ^h	ʃ ^h	g ^h
voiced affricates		dʒ		
aspirated clicks		ǀ ^h	ǀǀ ^h	ǀǀ ^h
prenasalized aspirated clicks		ǀǀ ^h	ǀǀǀ ^h	ǀǀǀ ^h
voiced fricatives	v	z	ʒ / ʒ	
breathy nasals	ᵐ	ᵑ	ᵑ̤	ᵑ̤̤
breathy glides	w̤		ɹ̤	
laryngeals	fi			

Lanham [1960:106] reports that, while both Zulu and the related Nguni language of Swati display tone displacement in the context of preceding depressors, in Xhosa, another Nguni tongue, the so-called “active measure” of tone displacement has been reported to be a mere tonal upglide: DV↓CV↓. Chuck Kisseberth [p.c. 1999] informs me that Xhosa speakers now freely vary between this tonal upglide, upglide-and-spread (DV↓CV↘), and displacement realizations (DV↓CV↘).

In perhaps the only invasive investigation of the depressor phenomenon, Traill et al. [1987] report that these consonants do not seem to possess the laryngeal posture characteristic of breathy phonation (that is, larynx lowering and vocal

fold vibration with an anterior chink). This does not surprise these researchers, since they do not hear breathy phonation. They do, however, observe fiber-optically that two of their three Zulu-speaking subjects show “an anterior movement of the arytenoids and the posterior movement of the tubercle of the epiglottis,” which, they claim “induces an unusual degree of shortening of the vocal chords [sic] with a consequent lowering of their rate of vibration” (p.265). They further suggest that some other speakers might lower pitch here with the better-known mechanism of cryothyroid laxing.

The earlier results of Rycroft [1980] are consistent with elements of Traill et al.’s findings, as he hears neither breathiness nor voicing on depressor stops. He does, however, report that the releases of these stops are breathy-voiced. In the context of a low tone, the following vowel is breathy for its entire duration. In the context of a high tone, the following vowel is breathy during its initial portion, and modal for its latter portion. Along with all other researchers, Rycroft reports that depressor consonants induce a far greater pitch lowering than do standard low tones. Finally, he suggests (p.17) that depressors might derive historically from fully breathy-voiced stops (i.e., those that are voiced on their closures as well as breathy upon their release), but have shifted over time to their current manifestation. This, notice, is consistent with their tonogenetic-like quality: as the voicing contrast here underwent attrition, the pitch perturbing effects that are a natural concomitant of voicing and breathiness may have become exaggerated (see Section 5 for further discussion). If Traill et al. are correct in their statement that epiglottal movement induces significant pitch lowering, this vocal tract configuration may be a latter-day reflex of the earlier closure-voicing distinction. Schachter [1976] posits a similar story for the historic origins of depressors in Swati, suggesting that the present-day “unnatural class” of depressors historically derives from voiced obstruents.

Beyond Nguni, depressor consonants have been observed in, for example, Tsonga [Traill & Jackson 1988], and Ikalanga, a Bantu language of the Shona group spoken in Botswana [Hyman & Mathangwane 1998, Mathangwane 1999]. These authors describe depressors as voiced obstruents, and observe that they block high tone spreading, convert high tones to rising contours, and induce an overall pitch lowering. Elsewhere, depressors have been observed in the Mijikenda Bantu subgroup, including Chikauma and Chirihe, which have been described in terms similar to those in Ikalanga, and engage in similar phonological behavior [Cassimjee & Kisseberth 1992]. Without a thoroughgoing instrumental investigation of all such patterns, however, the label “depressor” must suffice as a loose cover term for perhaps any number of phonetically distinct pitch-lowering and displacement-triggering consonants.

3. The sound substitution pattern: the phonology of Zulu tone displacement

The general phonological pattern of tone displacement consists of an apparent rightward shift of high-tones from short vowels when they are preceded by a consonant of the depressor class. The pattern is schematized and exemplified in (2). (Data from Cope [1966]; depressors are underlined; displaced tones are bold; depressor effects are italicized.) Spectrograms and pitch tracks (2b) show the first two forms, *i-si-t̤a:llo* and *i-z̤i-t̤a:llo*. Observe, especially, the difference in the pitch of the second vowels (indicated with upward arrows). The first form has a high tone, while the second form has a low tone, with the high being displaced. The high pitch at the beginning of the third syllable of *i-si-t̤a:llo* (indicated with the horizontal arrow) is probably a consequence of the preceding voiceless lateral, which is likely to induce a brief pitch increase at its offset due to high oral flow. The Zulu speaker is an adult male.

In (2), the presence of a depressor consonant occurs with a very low pitch—lower than phonological low tones—on the rightward vowel, and a high pitch beyond the release of the next consonant, after which the following tone is implemented. The pitch track clearly shows pitch changes that depressor consonants induce.

(2)a. schematic:

$$\begin{array}{c} \underline{D} \ V \ C \ V(:) \\ | \quad | \\ L \quad H \end{array}$$

examples:

i s i t̤ a:l o 'chair'
| | | |
L H L L

i z̤ i t̤ a:l o 'chairs'
| | ^ |
L L HL L

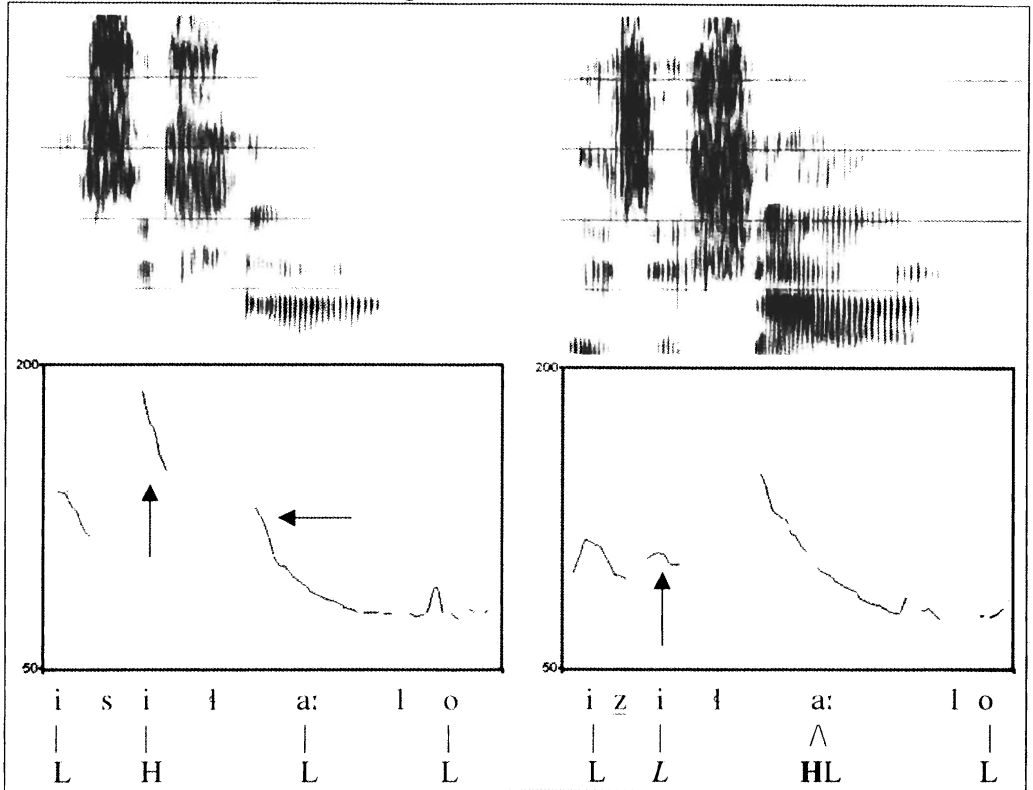
i[↓] n s i: z̤ w a 'young man'
| | |
H[↓] L L

j̤ i n s i: z̤ w a 'by a young man'
| ^ |
L HL L

i n̤ o: n̤ i 'bird'
| | |
H L L

n̤ e n̤ o: n̤ i 'with a bird'
| ^ |
L HL L

b. Sample spectrograms and pitch tracks



However, tone displacement does not take place under certain conditions. First, displacement does not take place from long vowels. In (3a) we observe no displacement on to a following syllable in the context of a long vowel, but displacement indeed takes place upon substitution with a short vowel. Phrase-penultimate vowels regularly lengthen, so displacement is not observed in this context either (3b). Furthermore, displacement does not take place from phrase-internal penults, which are not lengthened (3c). Finally, when the postvocalic consonant is also a depressor, displacement is not observed (3d).

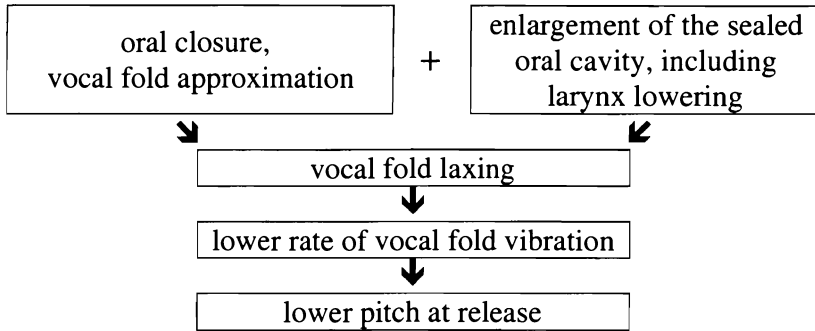
Cope reports that displacement is also blocked in the context of a following derived low-high toneme cluster “on analogy with [depressor-induced] rising tones” [1966:86]. Additionally, Lanham states that displacement is blocked “if the displaced toneme is to fall on a monosyllabic root” [1960:107].

those in Zulu may have their diachronic origins in certain phonetic forces on the sound system. I provide phonetic tutorials about pitch lowering at voiced (or breathy) stop releases, and about the speed at which speakers may produce pitch rises.

4.1 Laryngeal settings accompanying obstruents affect pitch at consonant release. It has often been noted that obstruents in general, and breathy-voiced obstruents especially, induce pitch perturbations upon their release [Hombert 1978, Maddieson 1978, Ohala 1978]. There are, in fact, reasonable aerodynamic and articulatory explanations for such effects. Both the pitch perturbation effects and one proposed explanation for them are briefly summarized here.

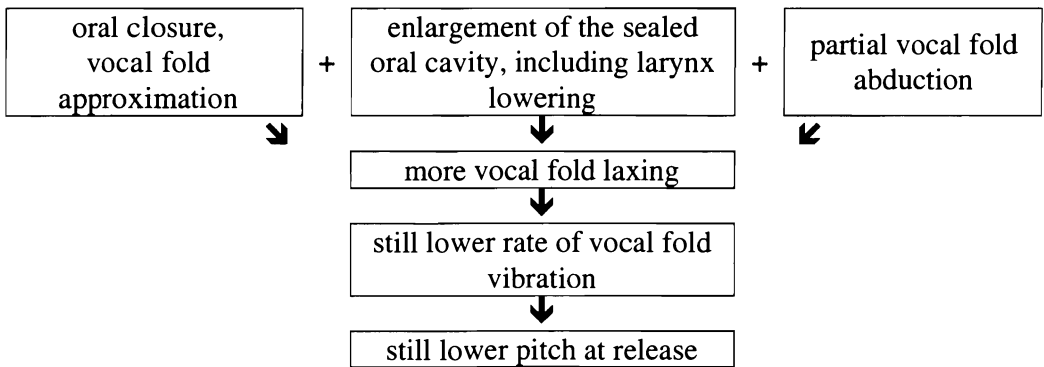
Voiced stops typically induce pitch lowering upon release, whereas voiceless stops typically induce pitch raising. Some researchers (for example, Ohala [1972], Ewan & Kronen [1974], Ewan [1976], Hombert, Ohala, & Ewan [1979]) implicate the tenseness versus laxness of the vocal folds at the point of stop release to account for these pitch perturbations. The route from contrastive voicing to vocal fold laxing is a fairly circuitous one, so allow me to elaborate by recapitulating the main thrust of these researchers' hypothesis. In certain contexts, contrastively voiced stops may be accompanied by an overall enlargement of the sealed supraglottal cavity (see, for example, Westbury & Keating [1986]). This cavity enlargement allows more air to cross the glottis into the sealed chamber, increasing the duration of transglottal flow, and increasing the duration of vocal fold vibration, which is likely to heighten the salience of closure voicing cues. This increased salience is important so that the voiced-voiceless contrast is maintained. Cavity expansion may be achieved in a number of ways, including cheek puffing, pharyngeal widening, velic raising, and larynx lowering. Larynx lowering may also involve an apparent automatic laxing of the vocal folds. It is this fold laxing which might contribute rather significantly to the observed pitch lowering observed at the release of voiced stops, as lax vocal folds are a well-attested concomitant of a lowered rate of vocal fold vibration, and lowered pitch [Hombert 1978]. This approach has been termed the "vertical tension hypothesis" by Hombert [1978]. Hombert, and also Riordan [1980] and Kingston [1985a,b], discuss certain empirical problems with the vertical tension hypothesis. However, all nonetheless concur that a positive correlation indeed exists between stop voicing and pitch lowering at release. The flowchart in (4) recapitulates the essentials of the vertical tension hypothesis. Indeed, the diachronic relationship between voiced stop release and pitch lowering is rather well established. In Yue Chinese, for example, an earlier voiced stop—voiceless stop contrast has evolved into low tone register—high tone register contrast [Karlsgren 1926].

(4) Flowchart of vertical tension hypothesis



When we add the partial vocal fold abduction necessary for breathy voicing, further pitch lowering may be induced, as the fold laxing which accompanies abduction typically occurs with a still-lower rate of vocal fold vibration (5). This has been observed in Hindi voiced aspirates, for example in Kagaya & Hirose [1975]. The diachrony of Punjabi is also consistent with this scenario. Here, voiced aspirates have evolved into plain stops followed by low tones [Gill & Gleason 1969, 1972].

(5) From breathy voice to pitch lowering



Breathy sonorants and voiced fricatives, like breathy voiced stops, possess both voicing and breath. Pitch lowering here has been observed as well.

Given the phonetic state of affairs during breathy phonation, pitch-lowering effects on the first portion of the following vowel in Zulu and elsewhere become readily understandable. And once again, recall that even if epiglottal movement is responsible for present-day pitch lowering here, this vocal tract configuration may

be seen as a latter-day reflex of an earlier closure-voicing distinction. The “trigger” of tone displacement may have changed, but the process remains.

The four patterns discussed—voiced stop, breathy-voiced stop, voiced fricative/breathy sonorant, and voiceless stop—plotted against their pitch-perturbing effects, are presented in schematic form in (6) (where H = higher pitch, L = lower pitch).

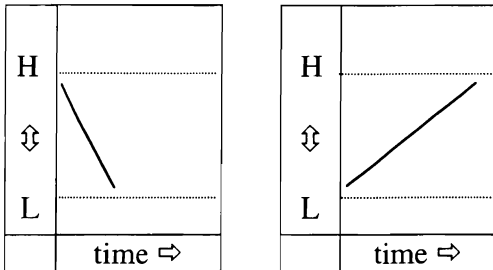
(6) Patterns of pitch perturbation

	d	a	d̤/ŋ̤/ð	a	t	a
H						
⇄	↗—		↗—		↘—	
L						

In fact, Löfqvist [1973] as well as Hombert [1975] show that, in some non-tonal languages, the pitch perturbing effects of the laryngeal settings of stops persist for more than 100 msec after release, far longer than a purely proximal aerodynamic account would predict; Ohala [1978:26] reports that the differing aerodynamic conditions present at obstruent release lasts “only a few milliseconds” beyond the beginning of voicing. It is this sort of finding—that phonetic explanations alone cannot fully account for language-specific production conventions, but nonetheless might serve to constrain the general direction of sound change—where functional forces on the system become relevant. I turn to this issue in Section 5. Before doing so, however, another relevant physiological finding must be presented.

4.2. Pitch rises take longer than pitch falls. The second phonetic finding that is relevant to tone displacement concerns the speed at which pitch contours can be induced by the laryngeal apparatus. Gandour [1974] reports that rising tones in Thai persist longer than do falling tones. Xu & Wang [2001] find a similar pattern in Mandarin Chinese. Moreover, in two studies [Ohala & Ewan 1973, Sundberg 1979] untrained subjects took a significantly greater amount of time moving from a lower-to-higher pitch across a given frequency range than vice versa. That is, pitch rises are accomplished more slowly than pitch falls of the same acoustic distance. A schematic is provided in (7).

(7) Time difference in pitch fall vs. pitch rise

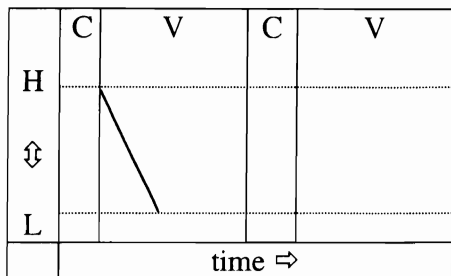


4.3. Tone displacement may be an illusion. These two distinct phonetic findings regarding pitch—pitch perturbation during breathy voice, and speed of pitch change—may have interacted in the diachrony of the Zulu tone displacement pattern. Specifically, when considering the state of the larynx in combination with the state of the supralaryngeal articulators, a characterization of the phenomenon as an active “displacement” may be viewed as illusory. Given the extra-low pitch following depressors (which, recall, is lower than the pitch of standard low tones), and also, given the sluggishness of pitch rises in comparison to pitch falls, the supralaryngeal articulators may already have achieved the proper configuration for a following consonant before the pitch rise is fully achieved. The result is the illusion of an active process of tone displacement: upon the release of this subsequent consonantal gesture, finally, the high pitch is achieved, and thus it seems to have flopped from one vowel to the next. As suggested by Ohala [1978:31], “[S]ince falling tones can be produced faster than rising tones...they might be less likely to ‘spill over’ onto the next syllable.” The schematic in (8d), which superimposes the relevant oral gestures on to the pitch contours, reveals the suggested illusion of tone displacement.

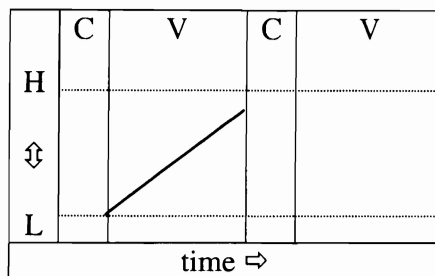
Note that we may now formulate a hypothesis about the apparent blocking of tone displacement from phrase-penultimate vowels. As these penults are automatically lengthened in Zulu, such vowels are apparently of sufficient duration to accommodate the pitch rise: the rise is achieved before the following consonant is implemented, and so there is no apparent displacement (9).

We are, in addition, now able to hypothesize about the blocking behavior of following depressors. As following depressors once again induce a pitch-lowering effect upon their release, there is equally little hope of the displaced high tone being realized in this context. Consequently, even in the context of a short vowel, such high tones are only realized at the end of their vowel of origin: here, the articulatory demands of the second depressor might indeed be characterized as “blocking” the propagation of the pitch rise. That is, in non-depressor contexts

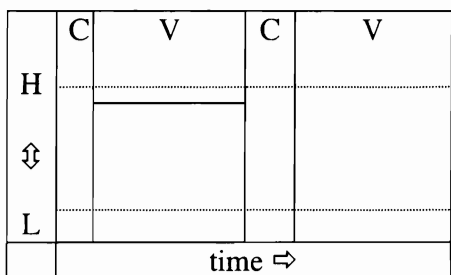
(8) a. HL sequence with no depressors



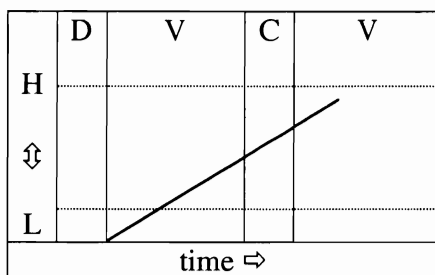
b. LH sequence with no depressors



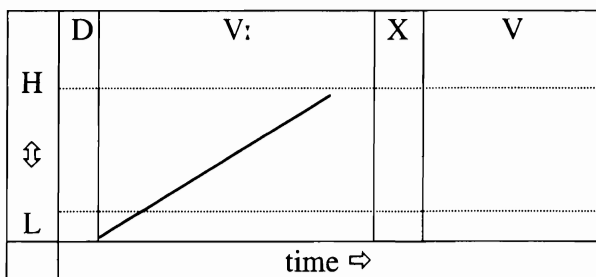
c. H with no depressor



d. H with depressor

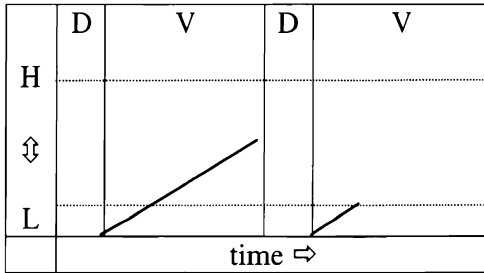


(9) Pitch rise without displacement



there are no conflicting articulatory demands made of the laryngeal apparatus that might inhibit the realization of the high tone. However, when a depressor intervenes between a high tone and the second vowel, the laryngeal musculature once again assumes the posture characteristic of depressors, thus inhibiting the high tone's realization on that second vowel (10). Thus, although depressors may no longer involve breathy phonation at their offset, the preceding scenario provides a working account of their paleophonetic origins.

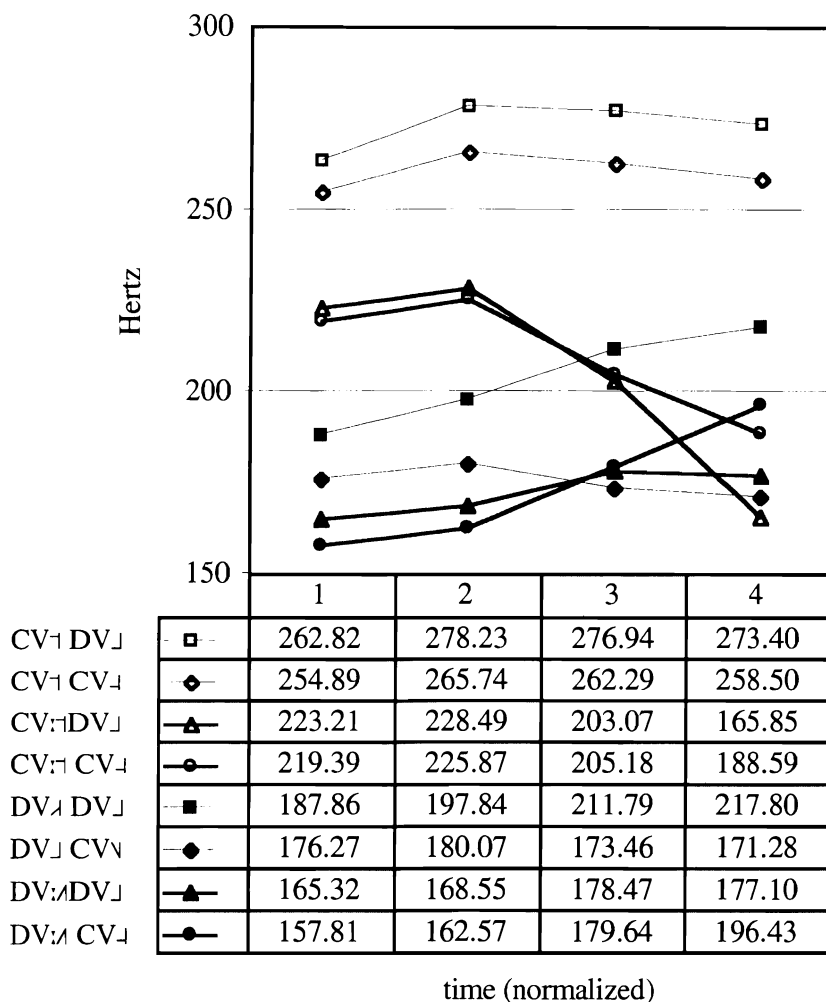
(10) H with intervening depressor

**4.4. Instrumental analyses support impressionistic descriptions of pitch effects.**

In a companion piece to the present study, Russell [2000] instrumentally investigates certain phonetic aspects of the interaction of tone, vowel length, and depressor consonants in Zulu. The figure in (11) provides pitch contour information for one of Russell's subjects, at four points along the duration of the initial vowel in the domain under investigation, pooled over eighteen tokens for each of her eight series (her second subject produced similar—though less pronounced—results). Point 1 is at consonant release, Point 2 is at the onset of a clear vowel formant structure, Point 3 is at vowel midpoint, and Point 4 is at vowel offset. The reader is cautioned that time is not drawn to relative scale: half the vowels are long (plotted with darker lines), and half are short (plotted with lighter lines).

Several aspects of this pitch patterning are relevant to the present discussion. Consistent with the impressionistic descriptions of all other researchers, Russell instrumentally confirms that depressor consonants co-occur with a much lower fundamental frequency at consonant release (solid points at time 1) than do non-depressors (hollow points at time 1). Short vowels preceded by non-depressors (CV₁ DV₁, CV₁ CV₁) begin at the highest frequencies, and retain this high range for their duration. Long vowels preceded by non-depressors (CV₁ DV₁, CV₁ CV₁) begin high (though not as high as their short counterparts), then lower over the course of the vowel. (In my experience, pitches which listeners label “high” are quite often falling in nature, rather than level high; this is especially true for long vowels). In at least one other study [Silverman et al. 1995], high tones on short vowels in Jalapa Mazatec are found to be higher in pitch than high tones on long vowels; the Zulu pattern matches the Jalapa Mazatec pattern in this respect. The short vowel preceded by a depressor and followed by a non-depressor (DV₁ CV₁) ends quite low. In this context, the high tone is displaced onto the following vowel. The origin of this pattern may be related to the hypothesis that there may be insufficient time for the pitch to significantly rise during the duration of the

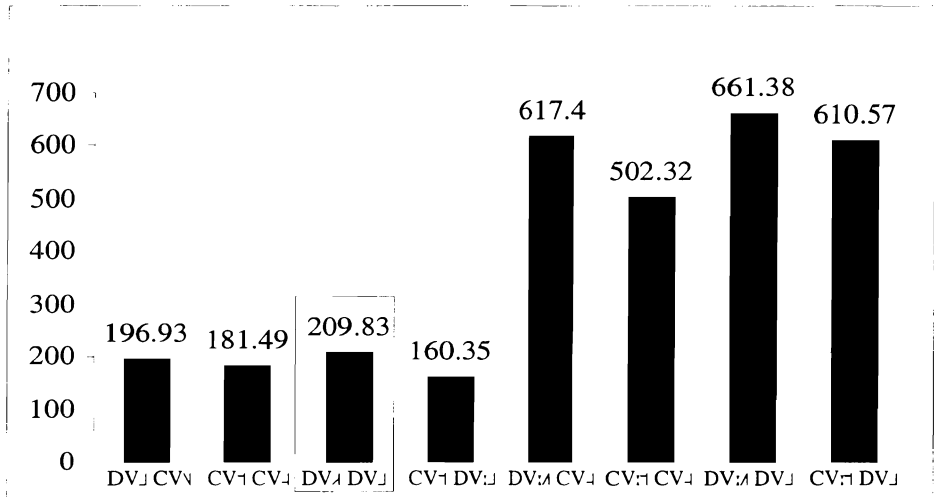
(11) Pitch changes (in Hz) over time for one speaker



short vowel. All other depressor-initial configurations (DV₁ DV₁, DV₁ DV₁, DV₁ CV₁) involve significant upglides. Here, both DV₁ DV₁ and DV₁ CV₁ are quite naturally expected to possess upglides over the course of their duration, since the vowel may be sufficiently long to accommodate a rise. The curious fact here is the presence of a rise on the DV₁ DV₁ configuration. All reports acknowledge the existence of a rise here. However, given a short vowel flanked by depressors, we might expect that a pitch rise here would be particularly meager. To begin to understand how this configuration accommodates the pitch rise,

consider the chart in (12), also from Russell [2000], which provides information about mean vowel length in all eight of the segmental contexts she considers.

(12) Mean vowel duration in msec for one speaker



With respect to short vowels, Russell determines that the vowel flanked by depressors (DV:] DV:] ; boxed) (for both her speakers) is significantly longer than the other short vowels (DV:] CV], CV:] CV], CV:] DV:]. While this longer vowel might be attributed to the automatic lengthening that typically accompanies a vowel which precedes a voiced consonant, this hypothesis does not seem sound given the extra-shortness of the CV:] DV:] configuration. Inspection of these same consonants with intervening long vowels does not support this hypothesis either, as, at least for this speaker, DV:] CV] is longer than CV:] DV].

Instead, the extra length in the context of DV:] D sequences may be seen as a functional accompaniment to the observed reduction in pitch height here: as pitch cannot be displaced in this context due to the presence of a following depressor, the vowel may be longer in order to provide a better opportunity for a pitch increase to be implemented, and consequently conveyed to listeners. Indeed, as Russell writes, "I attribute this extra length in the vowel in cell (c) [DV:] DV:] - D.S.] to the flanking effects of the surrounding depressor consonants...[W]ith no tone spread...to the following syllable, this vowel will be longer than in cases where depression need not be confined to the syllable proper." Interestingly, all syllable types with initial depressors are accompanied by longer vowels than their non-depressor-initial counterparts. Once again, this may be related to the slow pitch rise which is observed in these four contexts.

Russell further notes that long vowels flanked by depressors are longer than other long vowels, but that this is true only for one of her two speakers. Indeed, as Russell notes, “a long vowel need not...lengthen further to accommodate the effects of depression, as its extra length is sufficient...so that no changes need be made in vowel duration.”

To sum up, Russell instrumentally confirms that depressors induce significant pitch lowering at their offsets, and that displacement occurs when a following vowel is short and followed by a non-depressor. When the following vowel is short but followed by a depressor, no displacement is observed. However, the vowel here is significantly longer than other short vowels, thus better accommodating the pitch rise.

5. Tone displacement has functional value

I have thus far suggested that tone displacement in Zulu is in part understandable when placed in the context of (paleo-)phonetic aerodynamic and articulatory theories. At this point, speech scientists might protest that physiological limitations alone cannot be directly invoked to explain linguistic sound patterns: speakers exhibit exquisite control of their articulatory apparatus, and might modify their productions to accommodate to any raw, physical, ceiling- or floor-effects. For example, Hombert [1978] reports that pitch perturbations at voiced stop releases are significantly curtailed in tone languages such as Thai and Yoruba, whereas in certain non-tonal language (such as English), recall that these pitch perturbations may extend over 100 ms into the following vowel, well surpassing their perceptual limen, which Hombert determines to be around 40 ms. As Hombert notes, the English pattern cannot be accounted for by proximal phonetic forces alone, since aerodynamic theory does not predict such extensive pitch effects, and anyway, the “voiced” category is typically voiceless here [Lisker & Abramson 1964, Flege 1982]. However, it is reasonable to assume that speech communities arrive at their own consensus regarding the extent to which such phonetically natural states are exaggerated or curtailed. These conventions may surely be influenced by the language-specific system of contrasts in which these sound patterns find themselves (see especially Manuel [1999] on this point, who shows that language-specific patterns of coarticulation are influenced by language-specific systems of phonological contrast). For example, in the case of tonal languages such as Thai and Yoruba, Hombert notes that if stop-induced pitch perturbations went uncurtailed, the tonal system itself might suffer adverse functional consequences. By hypothesis then, such perturbations are curtailed here for this reason. In non-tonal languages such as English, however, there is no tonal system to encroach upon. Instead, extending these pitch perturbations well

into the following vowel may actually serve to enhance the contrast between prevocalic voiced and voiceless stops with no counter-functional consequences. Note that neither the Thai/Yoruba-type system nor the English-type system can be described as fully “natural” in that neither abides by purely physiological constraints with respect to pitch realization in this context.

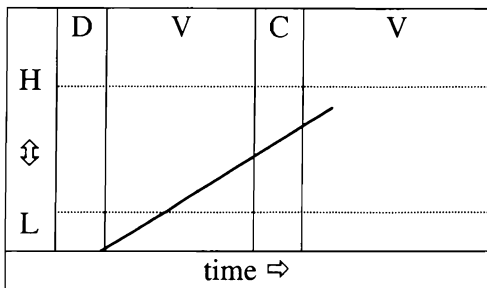
Returning now to tone displacement in Zulu, I argue in this section that functional factors do not in fact curtail the phonetic tendency toward tone displacement. Instead, they actually aid and abet it.

5.1. Non-displacement might neutralize contrasts. As discussed in Section 4, tone displacement is perhaps best regarded as a consequence of differences in the speed of implementation between oral and laryngeal gestures, which came to affect the diachrony of the system in the observed fashion. But merely because the sound substitution may be induced by a natural phonetic process does not entail that this process will strictly abide by exclusively physiological constraints, as the consideration of pitch-lowering effects of voiced obstruents has just demonstrated. Let us then imagine the functional consequences of tone displacement curtailment, that is, a situation in which high tones are not realized on a following vowel in the context of preceding depressors.

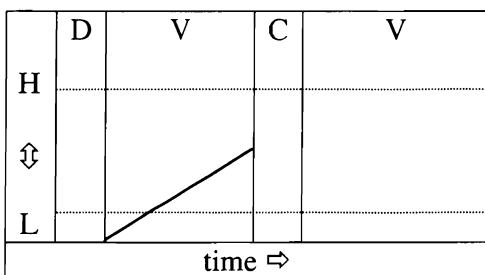
Consider the schematics in (13). In (13a), I repeat the figure from (8d) in which the higher pitch is achieved only as the following consonant is being released. In (13b), I present a hypothetical schematic involving the blocking of tone displacement in this same context.

What would be the functional consequences of blocking here? It is quite conceivable that the pitch rise, which is now limited only to the first vowel, would be insufficient for a reliable achievement of the high tone. Due to the only limited temporal domain in which the pitch rise is implemented, it may be sufficiently curtailed so that it might be misperceived by the listener as belonging to the low tone category. Indeed, recall that in the DVC context, pitch remains low for the duration of the vowel (DV↓ CV↓). This fact strongly suggests that past listeners reached exactly this conclusion: rising tones limited to short vowels which follow depressors may be misperceived as low tones. The functional consequence of such a misperception is loss of contrastive information, or neutralization. Indeed, deletion would result in two separate neutralizations, whereas displacement results only in a single neutralization. This is shown in (14). Here, both the deletion strategy and the displacement strategy are compared, along with unattested forms in which neither deletion nor displacement takes place (in quotation marks). It becomes clear that deletion halves the number of contrastive sequences, while displacement preserves three of four values.

(13) a. H with depressor



b. Hypothetical blocking of H displacement



(14) *unattested:*

“DV_↑ CV_↑”

“DV_↑ CV_↓”

“DV_↓ CV_↑”

“DV_↓ CV_↓”

deletion:

DV_↓ CV_↑

DV_↓ CV_↓

DV_↓ CV_↑

DV_↓ CV_↓

displacement:

DV_↓ CV_↑

DV_↓ CV_↓

DV_↓ CV_↑

DV_↓ CV_↓

Another possible scenario here involves the elongation of vowels when associated with pitch rises, as in Zulu DVD sequences. So, in DVC contexts, displacement is readily possible, and so the first vowel need not lengthen to accommodate the rise. By contrast, in DVD contexts, lengthening the vowel may serve to salvage the high tone contrast. As noted by Gandour [1977] (mentioned in Maddieson [1978]) this pattern is rather common at the cross-linguistic level.

Note that a functional approach may additionally account for the observed blocking of tone displacement in the context of a following low-high toneme cluster, for if displacement were to take place here, the rising contour would likely neutralize to high (15).

quently heard by listeners. In contrast, of course, phrase-internal final obstruents are sometimes followed by sonorants, which will not serve to inhibit the likelihood of voicing. In time, however, due in part to obstruent-final words alternating between utterance-final and utterance-internal position, learners may come to generalize the devoicing pattern to new contexts which bear decreasing phonetic and structural similarity to the original context, in that devoicing is less and less motivated by proximal phonetic forces—to the phrase, to the word, to the syllable. (See Hock [1998] for additional examples of this sort.) Thus, while the behavior of depressed tones in Zulu phrase-final penults may be phonetically and functionally motivated, we can begin to account for the expansion of this pattern into related domains that are not immediately amenable to phonetic explanation: non-displacement from phrase-final lengthened penults may have generalized to include all penults. In short, phrasal phonology may encroach upon lexical phonology.

6. Probability matching, and a proposed diachronic scenario

As stated in the previous section, it is well established that speakers exhibit exquisite control of their articulatory apparatus. Some researchers, for example Kingston & Diehl [1994], hypothesize that this prowess on the part of speakers is employed to improve the acoustic/auditory distinctness (and/or increase the articulatory ease) of phonological contrasts; that speakers exploit their phonetic knowledge to achieve functionally efficacious ends. In this section, I consider an alternative to this speaker-based approach to phonetic implementation and sound change, based primarily on the observation that token-to-token variability in speech production is largely recapitulated from generation to generation in a form of probability matching [Labov 1994, *pace* Gallistel 1990]. For example, if Zulu elders produce 95% of their tokens without displacement, and 5% with displacement, learners are likely to largely recapitulate these percentages in their own speech. Consequently, the inherent variation in speech production may not be as free as it is often thought to be, but may instead be conventionalized to a significant degree: the target of acquisition may be the variability itself. That is, the exquisite articulatory control that speakers display in their productions is best evidenced by the fact that they are able to largely match the variability present in the ambient pattern. The probability matching itself, no doubt, betrays an extremely sophisticated statistical analytic ability on the part of learners. Moreover, their actual productions betray evidence that they are able to implement their calculated probabilities in their own speech. On this view, learners' articulatory talents are harnessed in service to copying, not modifying (improving upon or otherwise) the ambient speech pattern. I suggest that a

listener-based account of the sound pattern offers a rather straightforward alternative to the speaker-based approach of theorists such as Kingston & Diehl.

Labov [1994] shows how probability matching may be affected by sound changes in progress. However, he does not comment on what might have induced such sound changes in the first place. In the case of Zulu at least, I have suggested that the origin of the sound change is phonetic. So, there may have been a point in the history of Zulu in which breathy-voiced consonants naturally induced an “unhinging” of following high tones such that the higher pitch fluctuated around the following consonant, being implemented with lower pitch when ceasing at the closure of the consonant, higher pitch when ceasing beyond its release, something like present-day Xhosa. This Xhosa-like variability may have been “pushed” toward a stabilized, displaced realization in the following manner: learners observe that there are, broadly speaking, two realizations of the high tone in this context, one in which the tone is displaced and higher in pitch, and one in which it is not displaced, and lower in pitch. These learners largely recapitulate the probability of each of the variants that are present in their parents’ (and, later, their peers’) speech. However, despite this general cross-generational stability in the probability of a particular variant being employed, there is nonetheless a phonetically natural tendency toward the displaced variant, and thus probabilities may yet slowly creep toward this value. This phonetically-induced creep may be aided and abetted by functional considerations: exactly because the displaced tokens result in a more accurate transmission of meaning—that is, there are fewer realizations that are neutralized upon displacement—the system may further creep in the direction of displacement. So, all displaced tokens may be perceived by listeners with the meanings that were intended by speakers. However, at least some of the non-displaced tokens will be miscategorized by listeners, as a consequence of neutralization. Consequently, the percentage of non-neutralized tokens (i.e., all the displaced tokens, but only most of the non-displaced tokens) should increase as these listeners become speakers: as some neutralized, non-displaced tokens are factored out of the total number of tokens for this category, the percentage of displaced tokens increases. The motivation for the creep towards this new convention should not be viewed as due to a teleological force: it is not a consequence of speakers wanting to be clearly understood by listeners. The stabilization is best viewed not as a cause, but rather as a consequence of effective communication. More specifically, I mean that those tokens in which tones were displaced resulted in a less ambiguous speech signal, one in which all contrasts (and all meanings) were recoverable. It is, by hypothesis, a consequence of the reliability of their perception—on the part of the listener—that high tones came to stabilize in the fashion they did, for listeners, having interpreted the

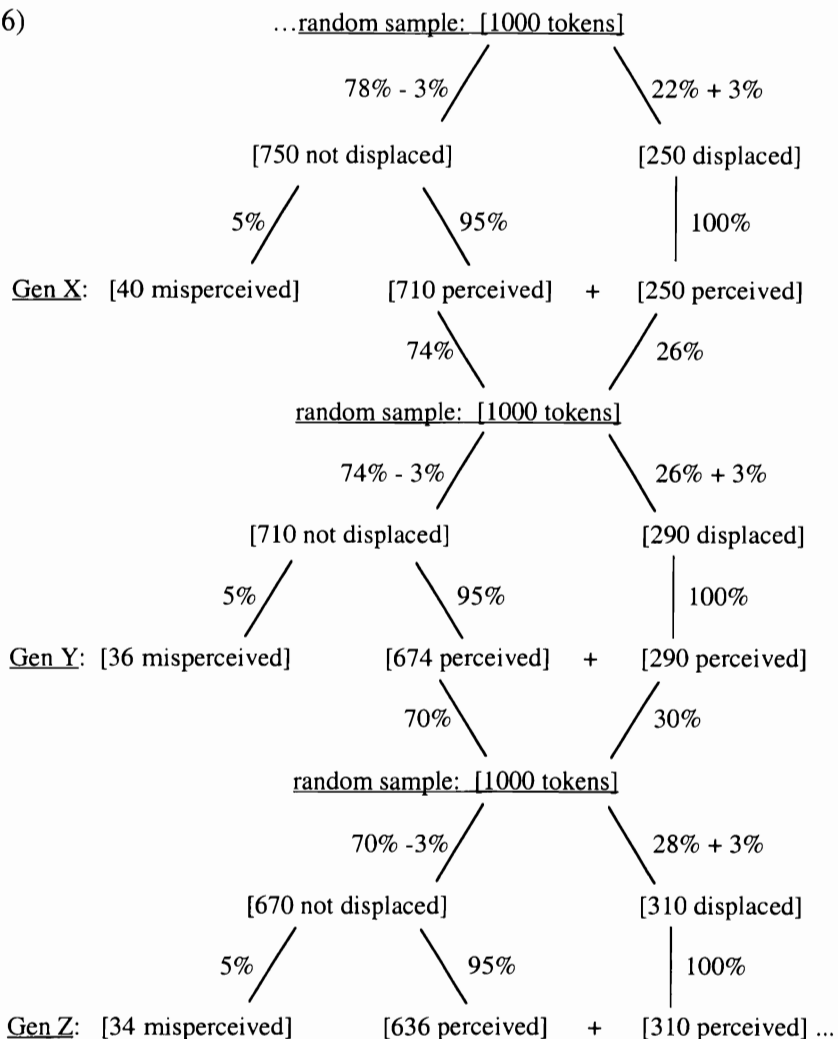
signal unambiguously, were far more likely to employ a similar articulatory routine when using these forms in their own speech.

We can now consider a schematic timetable, simplified to be sure, which captures the main forces argued to be at work in the diachrony of tone displacement. Consider the chart in 15. Entering the sound change midstream, we might take a 1000 token sample from Generation W. Of these tokens, 750 are non-displaced, while 250 are displaced. Most of these tokens are a consequence of learners' matching their probability of occurrence to the productions of Generation V. But given the natural tendency toward displacement, let's say that there is a 3% increase in the number of displaced tokens. In turn, Generation X perceives all displaced tokens unambiguously. Among the non-displaced tokens, however, let us suppose that a full 5% of these 750 tokens are misperceived by X due to the unlikelihood of clearly producing/perceiving the requisite pitch rise on only the first of the two vowels. These 40 misperceived tokens will not be pooled with those over which Gen X-ers calculate their probabilities.

Now we iterate the process: if we take a random sample of 1000 of Generation X's productions, we should observe that they match the probabilities that they perceive their elders to have produced, again allowing for a 3% phonetically-induced shift toward displacement. Generation X perceived 710 out of 960 tokens as non-displaced (40 tokens were misperceived); this constitutes a rate of 74%. Again subtracting 3% from this total, this results in 71%. So, out of 1000 tokens produced by Generation X, 710 will not be displaced, and 290 will be displaced. And again assuming that 5% of the non-displaced tokens will be misperceived by Generation Y, these children will perceive only 70% non-displaced tokens (674 of 964 tokens), and so on down the generations. We may now see, given the small tendency to misperceive non-displaced tokens, coupled with the small phonetic tendency toward an increased number of displaced tokens, how, over the course of time, the conventions of the system may undergo change.

There are, of course, any variety of different weights we might supply that would model different rates and directions of change, or model stability. Proper weighting of the proposed forces at work must of course be determined empirically. For example, in contexts and languages where we observe vowel lengthening (rather than tone displacement) to accommodate the pitch rise, those variants with increased vowel length would provide the germ of the sound change. Also, we have not yet considered the possibility that certain social trends might favor one realization over another, or any number of other conceivable variables, but a sufficiently sophisticated mathematical model would perhaps be able to approximately mirror such scenarios.

(16)



7. Dissimilar languages possess similar patterns

As is clear from the Hyman & Schuh report of 1974, the Zulu pattern of tone displacement is hardly unique: variations on the Zulu theme are found in a wide variety of African tonal languages and are also attested in languages bearing no areal or genetic relation to those discussed by Hyman & Schuh.

In Comaltepec Chinantec, an Otomanguan language of southern Mexico, high tones spread to a following syllable when immediately preceded by a tauto-syllabic low tone [Silverman 1997b]. The pattern is almost always allophonic,

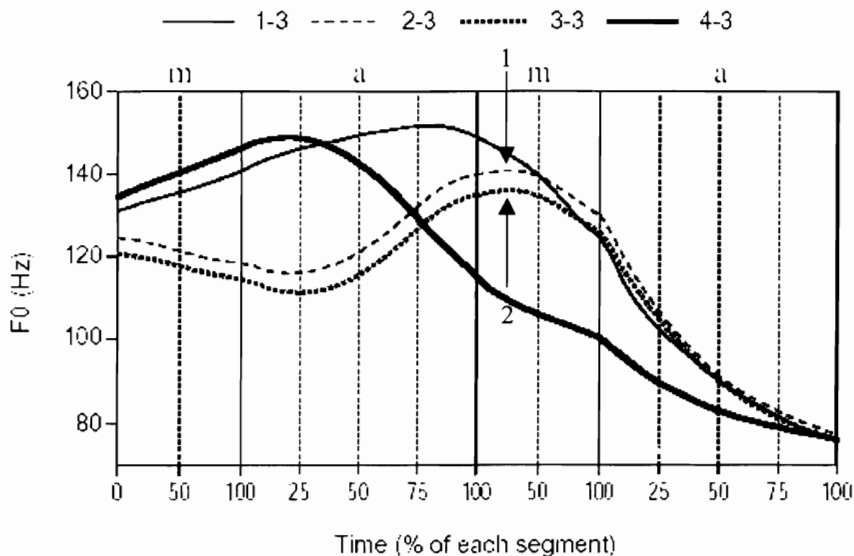
and only rarely neutralizing. For example, see the forms in (17) (from Pace [1990]).

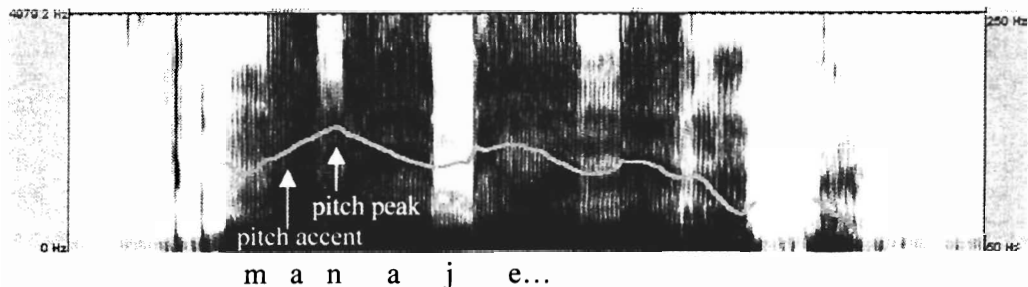
(17) Comaltepec Chinantec tone sandhi

non-sandhi context:	sandhi context:	gloss:
kwa ^Λ , to:⊥	kwa ^Λ to:∇	‘give a banana’
kwa ^Λ , ŋih ⊥	kwa ^Λ ŋih∇	‘give a chayote’
kwa ^Λ , ku:⊥	kwa ^Λ ku:∇	‘give money’
kwa ^Λ , ⁿ dʒu:⊥	kwa ^Λ ⁿ dʒu:∇	‘give a jug’

In Beijing Mandarin, tones with high offsets typically peak only after the following consonant has been implemented; tones with low offsets show a significantly lesser effect in these same contexts [Xu 1997, Xu & Wang 2001]. Figure (18), kindly provided by Yi Xu, shows four overlaid pitch tracks, showing various tone combinations during an all-sonorant bisyllabic sequence ([mama]). I call readers’ attention to the 2-3 pattern (rising-low; arrow 1) and 3-3 pattern (arrow 2,

(18) Mandarin Chinese [mama]



(19) Zagreb Croatian: *Manaje bila neznatna* (“A fault was insignificant”):

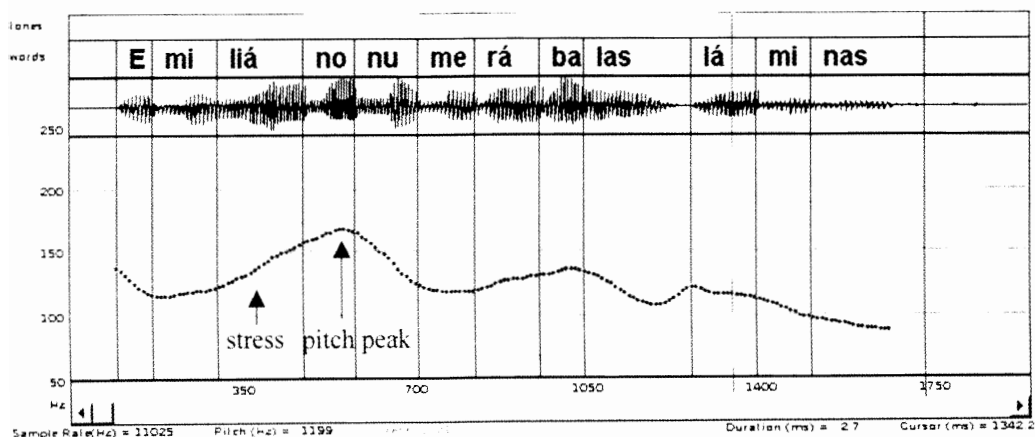
also rising-low, as a consequence of Mandarin tone sandhi). For both of these patterns, observe that the pitch peaks during the following nasal, not on the first vowel itself.

In Zagreb Croatian, high pitch-accented syllables possess a rising pitch contour, pitch peaks being realized on the post-tonic syllable, rather than on the accented syllable itself [Lehiste & Ivić 1986]. In (19) is a spectrogram, kindly provided by Rajka Smiljanić, showing a pitch accented syllable, with the pitch peak on the following nasal.

A similar pattern is observed in Peninsular Spanish: stressed syllables typically possess a pitch rise, with the pitch peak being realized on the post-stressed syllable [Navarro-Tomás 1944, Fant 1984, Prieto, van Santen, & Hirschberg 1995]. The pitch track in (20) was kindly provided by José Ignacio Hualde. The first word-stress here falls on the third syllable (“liá”). Pitch rises here, only to reach its maximum well into the following syllable (“no”).

(20) Peninsular Spanish:

Emiliano numeraba las láminas (“Emiliano was numbering the pictures”)



8. Concluding remarks

In summary, I have argued that physical properties of the speech mechanism—phonetic factors—may have initially induced the apparent “unhinging” of high tones in the context of preceding depressor consonants in Zulu. But independent functional factors may have induced the conventionalization of tone displacement. As tones were less likely to neutralize upon displacement, displaced tokens were less often ambiguously perceived, hence more likely to be reproduced.

More broadly, I hope to have shown that certain phonological patterns can be adequately accounted for by considering research results from a number of disparate fields of study: the state of any linguistic sound system is a consequence of phonetic, cognitive, and social forces that necessarily interact over time, passively shaped by the communicative function of language.

REFERENCES

- Beach, Douglas M. 1924. “The science of tonetics and its application to Bantu languages,” *Bantu Studies* 2: 75-106.
- Cassimjee, Farida and Charles W. Kisseberth. 1992. “The tonology of depressor consonants.” In Laura A. Buszard-Welcher, J. Evans, D. Peterson, L. Wee, and W. Weigel (eds.). *Special session on the typology of tone languages, Proceedings of the Berkeley Linguistics Society* 18: 26-40.
- Cope, A.T. 1960. “Zulu tonology.” *Afrika und Übersee* XLIII: 190-200.
- Cope, A.T. 1966. “Zulu phonology, tonology and grammar.” Ph.D. dissertation, University of Natal.
- Doke, Clement M. 1926. *The Phonetics of the Zulu Language*. Johannesburg: Witwatersrand University Press.

- Ewan, William G. 1976. "Laryngeal behavior in speech." Ph.D. dissertation, University of California at Berkeley.
- Ewan, William G. and Robert Krones. 1974. "Measuring larynx movement using the thyroumbrometer." *Journal of Phonetics* 2:327-335.
- Fant, L. 1984. *Estructura informativa en español. Estudio sintáctico y entonativo*. Acta Universitatis Upsaliensis 34.
- Fischer-Jørgensen, Eli. 1970. "Phonetic analyses of breathy (murmured) vowels in Gujarati." *Indian Linguistics* 28: 71-140.
- Flege, James Emil. 1982. "Laryngeal timing and phonation onset in utterance-initial English stops." *Journal of Phonetics* 10: 177-192.
- Fromkin, Victoria (ed.). 1978. *Tone: A Linguistic Survey*. New York: Academic Press.
- Gandour, Jack. 1974. "Consonant types and tone in Siamese." *Journal of Phonetics* 2: 337-350.
- Gandour, Jack. 1977. "On the interaction between tone and vowel length: Evidence from Thai dialects." *Phonetica* 34: 54-65.
- Gill, H.S. and H.A. Gleason. 1969. *A Reference Grammar of Punjabi*. Patiala.
- Gill, H.S. and H.A. Gleason. 1972. "The salient features of the Punjabi language." *Pakha Sanjam* 5: 1-50.
- Hock, Hans H. 1998. "Finality, prosody, and change." MS. University of Illinois at Urbana-Champaign.
- Hombert Jean-Marie. 1975. "Towards a theory of tonogenesis: an empirical, physiologically and perceptually-based account of the development of tonal contrasts in languages." Ph.D. dissertation, University of California, Berkeley.
- Hombert Jean-Marie. 1978. "Consonant types, vowel quality, and tone." In V. Fromkin (ed.), *Tone: A Linguistic Survey*. New York: Academic Press. Pp. 77-111.

- Hombert Jean-Marie, John J. Ohala, and William G. Ewan. 1979. "Phonetic explanations for the development of tones." *Language* 55: 37-58.
- Hyman, Larry M. and Joyce T. Mathangwane. 1998. "Tonal domains and depressor consonants in Ikalanga." In L. M. Hyman and C. W. Kisseberth (eds.), *Theoretical Aspects of Bantu Tone*. Stanford: CSLI Publications.
- Hyman, Larry M. and Russell G. Schuh. 1974. "Universals of tone rules: evidence from West Africa." *Linguistic Inquiry* 5.1: 81-115.
- Kagaya, Ryohei and Hajime Hirose. 1975. "Fiberoptic, electromyographic, and acoustic analyses of Hindi stop consonants." *Annual Bulletin*, Research Institute of Logopedics and Phoniatics, Tokyo 9: 27-46.
- Karlgren, Bernard. 1926. *Études sur la phonologie chinoise* (Archives d'Études Orientales 15). Stockholm: Norstedt.
- Kingston, John. 1985a. "The phonetics and phonology of the timing of oral and glottal events." Ph.D. dissertation, University of California, Berkeley.
- Kingston, John. 1985b. "The linguistic use of vertical larynx movement." MS., University of Texas at Austin.
- Kingston, John and Randy L. Diehl. 1994. "Phonetic knowledge." *Language* 70: 419-454.
- Khumalo, J.S.M. 1981. "Zulu tonology, part 1." *African Studies* 40: 53-130.
- Labov, William S. 1994. *Principles of Linguistic Change*. Oxford: Blackwell.
- Lanham, L. W. 1958. "The tonemes of Xhosa." *African Studies* 17: 65-81.
- Lanham, L. W. 1960. "The comparative phonology of Nguni." Ph.D. dissertation, Witwatersrand University.
- Lehiste, Ilse and P. Ivić. 1986. *Word and Sentence Prosody in Serbocroatian*. Cambridge: MIT Press.
- Lisker, Leigh and Arthur Abramson. 1964. "A cross-linguistic study of voicing in initial stops: acoustical measurements." *Word* 20: 384-422.

- Löfqvist, Anders. 1973. "On intrinsic and extrinsic F0 variation in Swedish tonal accents." *Monthly Internal Memorandum*. March: 19-32.
- Maddieson, Ian. 1978. "Universals of tone." In J. H. Greenberg (ed.), *Universals of Human Language*. Stanford: Stanford University Press. Pp. 335-365.
- Maddieson, Ian. 1997. "Phonetic universals." In William J. Hardcastle and John Laver (eds.), *The Handbook of Phonetic Sciences*. Oxford: Blackwell. Pp. 619-639.
- Manuel, Sharon. 1999. "Cross-language studies: relating language-particular co-articulation patterns to other language-particular facts." In William J. Hardcastle and Nigel Hewlett (eds.), *Coarticulation*. Cambridge: Cambridge University Press. Pp. 179-199.
- Mathangwane, Joyce T. 1999. *Ikalanga Phonetics and Phonology: A Synchronic and Diachronic Study*. Stanford: CSLI Publications.
- Navarro-Tomás, Tomás. 1944. *Manuel de entonación española*. New York: Hispanic Institute in the United States.
- Ohala, John J. 1978. "Production of tone." In V. Fromkin (ed.), *Tone: A Linguistic Survey*. New York: Academic Press. Pp. 5-39.
- Ohala, John J. and William G. Ewan. 1973. "Speed of Pitch Change." (Abstract) *Journal of the Acoustical Society of America* 53: 345.
- Pace, Wanda Jane. 1990. "Comaltepec Chinantec Verb Inflection." In W. R. Merrifield and C. R. Rensch (eds.), *Syllables, Tone, and Verb Paradigms*. (Studies in Chinantec Languages, vol. 4). Summer Institute of Linguistics. Pp. 21-62.
- Prieto, Pilar, Jan van Santen, and Julia Hirschberg. 1995. "Tonal alignment patterns in Spanish." *Journal of Phonetics* 23: 429-451.
- Riordan, C. J. 1980. "Larynx height during English stop consonants." *Journal of Phonetics* 8: 353-360.

- Russell, Margaret. 2000. "Phonetic aspects of tone displacement in Zulu." In Arika Okrent and John Boyle (eds.), *The Proceedings of the Chicago Linguistic Society* 36.1: 427-439. Chicago: Chicago Linguistic Society.
- Rycroft David K. 1980. "The depression feature in Nguni languages and its interaction with tone." Grahamstown: Rhodes University, Communication no. 8, Department of African Languages.
- Schachter, Paul. 1976. "An unnatural class of consonants." in L. M. Hyman, L. C. Jacobson, and R. G. Schuh (eds.), *Papers in African Linguistics in Honor of William E. Welmers. Studies in African Linguistics, Supplement 6*: 211-220.
- Silverman, Daniel. 1997a. "Laryngeal complexity in Otomangean vowels." *Phonology* 14.2: 235-261.
- Silverman, Daniel. 1997b. "Tone sandhi in Comaltepec Chinantec." *Language* 73.3: 473-492.
- Silverman, Daniel, Barbara Blankenship, Paul Kirk, and Peter Ladefoged. 1995. "Phonetic structures in Jalapa Mazatec." *Anthropological Linguistics* 37.1: 70-88.
- Sundberg, Johan. 1979. "Maximum speed of pitch changes in singers and untrained subjects." *Journal of Phonetics* 7.2: 71-79.
- Traill, Anthony and Michel Jackson. 1988. "Speaker variation and phonation type in Tsonga." *Journal of Phonetics* 16: 385-400.
- Traill, Anthony, J.S.M. Khumalo, and P. Fridjhon. 1987. "Depressing facts about Zulu." *African Studies* 46: 255-274.
- Tucker, Archibald N. 1949. "Sotho-Nguni orthography and tone-marking." *Bulletin of the School of Oriental and African Studies* 13.1: 200-224.
- Westbury, John R. and Patricia A. Keating. 1986. "On the naturalness of stop consonant voicing." *Journal of Linguistics* 22: 145-166.
- Xu, Yi. 1997. "Contextual tonal variation in Mandarin." *Journal of Phonetics* 25: 61-83.

Xu, Yi and Q. Emily Wang. 2001. "Pitch targets and their realization: evidence from Mandarin Chinese." *Speech Communication* 33: 319-337.

Department of Linguistics
4088 Foreign Languages Building
707 S. Mathews
University of Illinois
Urbana, IL 61801
dan@uiuc.edu

[Received December 2000;
accepted June 2001]